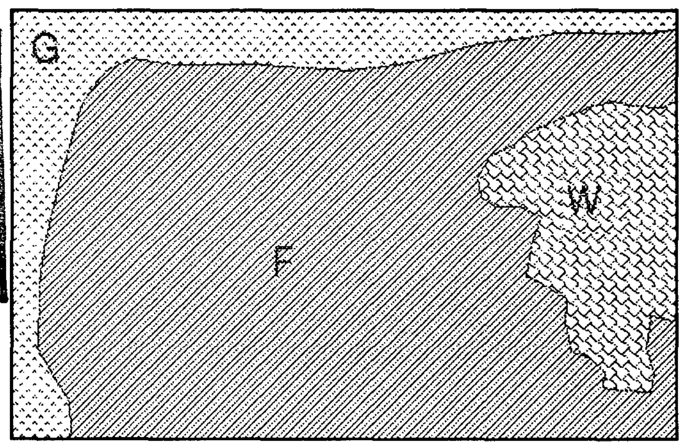
US Army Corps
of Engineers
Construction Engineering
Research Laboratories

USACERL Technical Report EN-95/04
April 1995

Accuracy Assessment of the Discrete Classification of Remotely-Sensed Digital Data for Landcover Mapping

by
Gary M. Senseman, Calvin F. Bagley, and Scott A. Tweddale



Remotely-sensed digital data may potentially help natural resource managers of military installations derive landcover information needed to inventory and monitor the condition of publicly owned lands. One method of deriving landcover information is to perform a discrete classification of remotely-sensed digital data. Before using a remote-sensing derived landcover map in management decisions, however, an accuracy assessment must be performed.

This study compared methods of site-specific and non-site-specific accuracy assessment analyses in the context of deriving a general landcover map. Non-site-specific analysis was found to be useful only for detecting gross errors in a classification. Site-specific analysis was found to provide critical information about a classification's locational accuracy. The use of an error matrix was also found to provide additional insight into classification errors, and the use of the Kappa Coefficient of Agreement was found to account for random chance in the accuracy assessment. At a minimum, a Kappa Coefficient of Agreement should be attached to any resultant classification of satellite imagery. Ideally, several measure of accuracy assessment should be performed and included as documentation with any classification.

USER EVALUATION OF REPORT

REFERENCE: USACERL Technical Report EN-95/04, *Accuracy Assessment of the Discrete Classification of Remotely-Sensed Digital Data for Landcover Mapping*

Please take a few minutes to answer the questions below, tear out this sheet, and return it to USACERL. As user of this report, your customer comments will provide USACERL with information essential for improving future reports.

1. Does this report satisfy a need? (Comment on purpose, related project, or other area of interest for which report will be used.)

---

2. How, specifically, is the report being used? (Information source, design data or procedure, management procedure, source of ideas, etc.)

---

3. Has the information in this report led to any quantitative savings as far as manhours/contract dollars saved, operating costs avoided, efficiencies achieved, etc.? If so, please elaborate.

---

4. What is your evaluation of this report in the following areas?

a. Presentation: ______

b. Completeness: ______

c. Easy to Understand: ______

d. Easy to Implement: ______

e. Adequate Reference Material: ______

f. Relates to Area of Interest: ______

g. Did the report meet your expectations? ______

h. Does the report raise unanswered questions? ______

i. General Comments. (Indicate what you think should be changed to make this report and future reports of this type more responsive to your needs, more usable, improve readability, etc.)

5. If you would like to be contacted by the personnel who prepared this report to raise specific questions or discuss the topic, please fill in the following information.

Name: ______________________

Telephone Number: ______________________

Organization Address: ______________________

____

____

6. Please mail the completed form to:

Department of the Army
CONSTRUCTION ENGINEERING RESEARCH LABORATORIES
ATTN: CECER-IMT
P.O. Box 9005
Champaign, IL 61826-9005

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave Blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	April 1995	Final

4. TITLE AND SUBTITLE
Accuracy Assessment of the Discrete Classification of Remotely-Sensed Digital Data for Landcover Mapping

5. FUNDING NUMBERS
4A162720
A896
NN-TS4

6. AUTHOR(S)
Gary M. Senseman, Calvin F. Bagley, and Scott A. Tweddale

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Army Construction Engineering Research Laboratories (USACERL)
P.O. Box 9005
Champaign, IL 61826-9005

8. PERFORMING ORGANIZATION REPORT NUMBER
EN-95/04

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Center for Public Works (USACPW)
ATTN: CECPW-FN
7701 Telegraph Road
Alexandria, VA 22312-3862

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES
Copies are available from the National Technical Information Service, 5285 Port Royal Road, Springfield, VA 22161.

12a. DISTRIBUTION / AVAILABILITY STATEMENT	12b. DISTRIBUTION CODE
Approved for public release; distribution is unlimited.	

13. ABSTRACT (Maximum 200 words)

Remotely-sensed digital data may potentially help natural resource managers of military installations derive landcover information needed to inventory and monitor the condition of publicly owned lands. One method of deriving landcover information is to perform a discrete classification of remotely-sensed digital data. Before using a remote-sensing derived landcover map in management decisions, however, an accuracy assessment must be performed.

This study compared methods of site-specific and non-site-specific accuracy assessment analyses in the context of deriving a general landcover map. Non-site-specific analysis was found to be useful only for detecting gross errors in a classification. Site-specific analysis was found to provide critical information about a classification's locational accuracy. The use of an error matrix was also found to provide additional insight into classification errors, and the use of the Kappa Coefficient of Agreement was found to account for random chance in the accuracy assessment. At a minimum, a Kappa Coefficient of Agreement should be attached to any resultant classification of satellite imagery. Ideally, several measure of accuracy assessment should be performed and included as documentation with any classification.

14. SUBJECT TERMS		15. NUMBER OF PAGES
Land Management Remote Sensing Digital Data	Geographic Information Systems	30
		16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	SAR

Foreword

This study was conducted for U.S. Army Center for Public Works (USACPW) under Project 4A162720A896, "Environmental Quality Technology"; Work Unit NN-TS4, "Imagery Data for Training Area Management." The technical monitor was Victor Diersing, CECPW-FN.

The work was performed by the Environmental Natural Resources Division (EN) of the Environmental Sustainment Laboratory (EL), U.S. Army Construction Engineering Research Laboratories (USACERL). Dr. William Severinghaus is Chief, CECER-EN, and William Goran is Chief, CECER-EL. The USACERL technical editor was William J. Wolfe, Information Management Office.

LTC David J. Rehbein is Commander and Acting Director of USACERL, and Dr. Michael J. O'Connor is Technical Director.

Contents

1 Introduction

Background

The U.S. Army is responsible for managing 12.4 million acres of land on 186 major installations worldwide (DA 1989). Army installations span all North American ecoregions and habitat types, and are used for a variety of military training and testing activities, along with many nonmilitary uses, including fish and wildlife, forest products, recreation, agriculture, and grazing. Proper land management supports the military mission and multiple use activities, but also presents the Army with a unique challenge as monitor and steward of public lands. The Army's standard for land inventory and monitoring is the Land Condition-Trend Analysis (LCTA) program, which it uses to collect, analyze, and report natural resources data.

Programs like LCTA depend on accurate data—e.g., information on the composition and distribution of landcover—to help Army land managers maintain installation natural resources. Such information may be compiled from various data sources and presented in many forms. Landcover characteristics, for example, may be reported in summary statistics or spatial representations. An increasingly common approach is to represent landcover information in spatial form derived from remotely-sensed multispectral satellite digital data. A geographic information system (GIS) with image-processing capabilities can help automate the process of interpreting this digital data to identify and delineate various characteristics of the earth's surface.

For information derived from remotely-sensed data to be useful in decisionmaking, the data must be checked against the physical land features for accuracy; it must be "accuracy assessed." Much attention has been given to the various methods of classifying remotely-sensed digital data; however, less attention has been paid to the rigorous accuracy assessment of the classification products. Product maps are often prematurely presented as successful integration of ground-truthed data without any statistical evaluation or with only a weak application of inappropriate statistical methods. An objective, quantitative, statistical approach is required to estimate the accuracy of thematic classifications of remotely-sensed digital data. Adequate techniques of classification accuracy have been developed and must now be applied to verify data collected for use in applications developed to support natural resource management through the LCTA program (ETN 420-74-3 1990; Tazik et al. 1992).

Objectives

The objectives of this study were to:

1. Review applicable methods to perform accuracy assessment of remotely sensed data when used for natural resources inventory and monitoring goals
2. Derive a framework for use of applicable methods within the LCTA program
3. Test the execution of accuracy assessment methods with synthetic data.

Approach

A literature search was done to review current accuracy assessment methods. Current methods were located, characterized, and presented in detail for the user of LCTA data. Procedures were derived for performing accuracy assessment, and a test example was performed using synthetic data.

Mode of Technology Transfer

It is recommended that methods for accuracy assessment outlined in this report be incorporated into the LCTA program, and that these or similar methods be required for all applications of remotely sensed data when product map layers are to be used in natural resources management decisionmaking.

2 Remotely-Sensed Satellite Digital Data

Two types of data are required to complete any sort of mapping with remotely-sensed data. The first is the remotely-sensed data itself, and the second (and equally important) is the *ground-truthed data*. Without ground-truthed data, remotely-sensed data is of limited value.

Remotely-sensed data is ". . . acquired by a device that is not in contact with the object, area, or phenomenon under investigation" (Lillesand and Kiefer 1987). Remotely-sensed data can be acquired by space-based, airborne, or ground-based electromagnetic sensors. These devices measure the electromagnetic energy being reflected from earth for a particular area. This electromagnetic energy is an ordered array of radiation extending from short to long radio waves. Remote sensor systems separate these radio waves into distinct bands or channels, analogous to the colors of the spectrum (Jensen 1992, pp 32-36). The commercial French satellite SPOT (*Systeme Probatoire pour l'Observation del la Terre*) and LANDSAT, a U.S.-developed series of platforms operated by Earth Observation Satellite Company (EOSAT) are the most common satellite systems for large area data acquisition. Each of these satellites sensors offers a broad area of coverage. The SPOT high resolution visible (HRV) sensor covers 60 × 60 km with a spatial resolution of 20 m, temporal resolution of 26 days and three spectral bands (green, 0.50 to 0.59 μm; red, 0.61 to 0.68 μm; and near infrared, 0.79 to 0.89 μm). The LANDSAT Thematic Mapper (TM) sensor covers 165 × 180 km with a temporal resolution of 16 days, six bands with spatial resolution of 30 m (blue, 0.45 to 0.52 μm; green, 0.52 to 0.60 μm; red, 0.63 to 0.69 μm; near infrared, 0.76 to 0.90 μm; mid infrared, 1.55 to 1.75 μm; and mid infrared, 2.08 to 2.35 μm) and one band with spatial resolution of 120 m (thermal infrared, 10.4 12.5 μm). Both systems are in Sun-synchronous orbits so the satellite passes over the same area of the earth at the same solar time in each temporal cycle.

Classification

Multi-spectral, remotely-sensed digital data can provide a great deal of information on characteristics of the Earth's surface. Various image-processing techniques, when applied to this data, enhance the extraction of earth resource information. Two basic forms of derived information are continuous and discrete data. Continuous data can

be represented as a continuum such as percent bare ground or cover. Discrete data is separate and distinct, or discontinuous. Examples of thematic representations of discrete variables are soils, plant communities, and land use.

Remotely-sensed spectral data is a continuous form of data where digital numbers (dn) represent the reflected energy in each band (spectral region). The objective of image classification is to assign each cell or picture element (pixel) of the satellite digital data into an appropriate thematic category in a process called "discrete classification." The most common data clustering algorithm used to automate classification of satellite digital data is *maximum likelihood*. Other types of data clustering algorithms are the *minimum distance*, *mahanbois distance*, and *contextual* (smap).

Accuracy Assessment

For remotely-sensed data to be truly useful and effective, an appropriate technique of accuracy assessment needs to be performed. Accuracy assessment can be defined as a comparison of a map produced from remotely-sensed data with another map from some other source. A determination is made of how closely the new map produced from the remotely-sensed data matches the source map. Evaluation of the accuracy of a classification of remotely-sensed data can fall into one of two general categories: non-site-specific assessment, or site-specific assessment (Campbell 1987). Of several approaches to accuracy assessment, the following sections will focus on the site-specific error analysis of pixel misclassification.

Non-Site-Specific

Non-site-specific assessment is a simplistic approach to assessing the accuracy of the classification of remotely-sensed data (Campbell 1987, p 340). In this method, a comparison is made between the "known" or estimated area and the area derived through the process of the discrete classification of remotely-sensed data. For example, an estimate is made of the percentage of area represented by three categories: grassland, woodland, and water. Suppose these "known" map areas are estimated to be 20 percent grassland, 60 percent forest, and 20 percent water. It is then possible to compare the estimated area by category to the classified imagery-derived areas for each category. The areas for each category derived from the discrete classification of the remotely-sensed data consist of: 18 percent grassland, 61 percent forest, and 21 percent water. After classification of the remotely-sensed data, non-sites error assessment of the derived map is done. Assuming that the area estimates of each of the three categories are correct, the non-site-specific error analysis would not indicate a significant problem with the classification of the remotely-sensed data.

Non-site-specific error analysis consists of identifying general problems with the resulting classification, but provides no information about the locational accuracy of the assessment (pixel misclassification), or how well each pixel was classified. Even though there was close agreement between the estimated areas and the areas derived from the classified map, the classification may still have been *inaccurate* in terms of locational or site-specific errors. If a substantial difference between the total areas in the estimate and the total areas in the classification occurred, it would be clear that the classification had not performed well. Thus, limitations of using non-site-specific error assessment quickly reveal themselves. Figures 1 and 2 show the limitation of non-site-specific error assessment for discrete classification relative to locational errors. Figure 1 shows the "known" data themes and is the reference map. Figure 2 is the result of the discrete classification of the remotely-sensed data set. The proportions of the three categories are similar in each map, but the physical location of each category in the resultant map (Figure 2) does not match the original map (Figure 1).

Non-site-specific accuracy assessment has limited utility; it is useful only for detecting gross problems with discrete classifications because of its inherent inability to identify locational errors. In other words, non-site-specific accuracy assessment can provide some measure of agreement between a reference map and classification in terms of the areal extent of each category, but it does not provide any information about the locational accuracy of the classification. Locational accuracy is important if the objective is to derive some form of spatial representation of landcover characteristics from the classification of remotely-sensed data. Results derived from an error assessment using the non-site-specific technique may be misleading. Site-specific error analysis is a more rigorous technique for assessing accuracy.

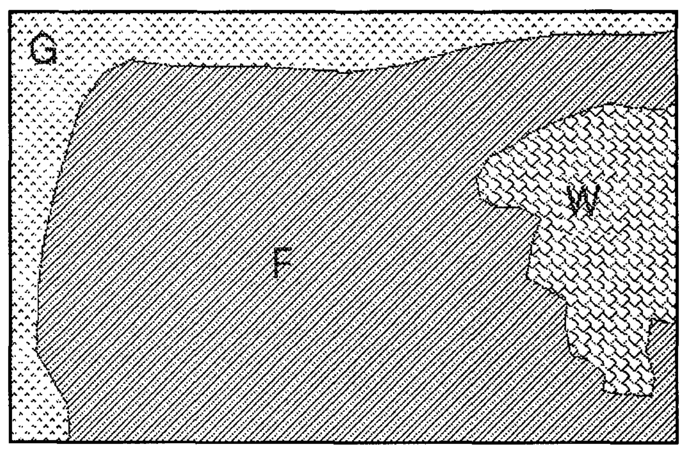


Figure 1. Reference map (G=Grass, F=Forest, W=Water).

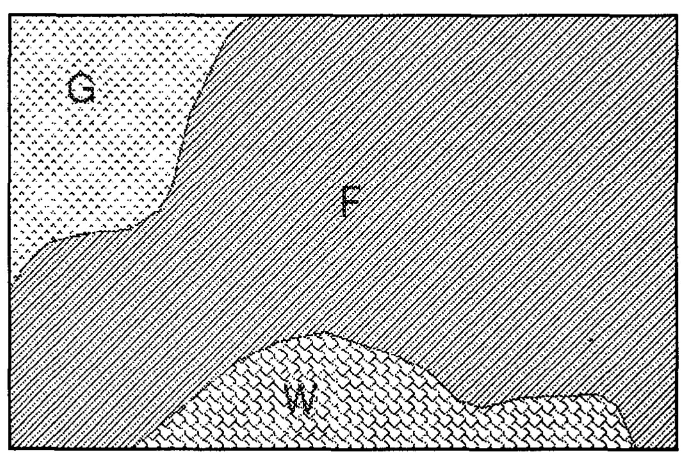


Figure 2. Classification map (categories same as Figure 1).

Site-Specific

Site-specific error analysis takes into account locational accuracy of the classification (Campbell 1987, p 340). This process makes a pixel-by-pixel comparison between the remotely-sensed, data-derived thematic map and a "true" map of the area with the same theme. This accuracy assessment approach is still prone to errors attributable to control point location error, boundary line error, and pixel misclassification (Hord and Brooner 1976). Usually, the purpose of classification is to derive a thematic map of some unknown characteristic of the Earth's surface or some characteristic that has changed over time, so it would be unusual for a complete and *current* reference map to exist. However, the reference map can be represented by a sample of locations within each theme for the area of interest. The selection sample locations and sample size is determined by the requirements of the subsequent analysis. In most cases, the analysis will include inter-class analysis as well as overall accuracy analysis.

Data requirements, sampling approach, and sample size. The data requirements for performing a classification include remotely-sensed data, ground-truthed training data for characterizing spectral parameters of each class (e.g., "plant community type"), and an independent set of ground-truthed data (reference data) for accuracy assessment. Since it is impractical to have a complete pixel-by-pixel "ground truth" map, an adequate subset or sample number of points (pixels) is needed for there to be a rigorous accuracy assessment of a classification. One must use an appropriate sampling technique that meets statistical requirements.

Site-specific accuracy assessment can be evaluated for an overall classification or on a per-category basis. The more rigorous and useful approach is to evaluate accuracy on a per-category basis, which provides more insight into classification errors that may be unique to specific categories. Category specific errors are not as readily apparent in an overall assessment. A stratified random method is an appropriate sampling method for accuracy assessment on a per-category basis (Van Genderen and Lock 1977). The Kappa Coefficient of Agreement, which is a statistical measure of the significance of difference between observed agreement of two classifications versus agreement due to random chance, is commonly used in both types of assessment and requires a multinomial sampling method. A stratified random sample is a multinomial sampling method, and therefore is an appropriate sampling method to be used with the Kappa statistic. The Kappa statistic is discussed in more detail later in this document. With the stratified random approach, points are stratified by map category, and simple random sampling is employed within each stratum (Stehman 1992). Once the sampling design has been determined, the number of sample points must be determined. The number of reference pixels required for accuracy assessment depends on the minimum level of accuracy (e.g., 85 percent) required. Jensen discusses

equations suitable for determining a minimum number of pixels required for different levels of accuracy (Jensen 1986). One approach to determining the total number of reference pixels (observations) needed to assess the accuracy at a minimum level uses Equation 1:

$$N = \frac{4(p)(q^{\sim})}{E^2} \qquad \text{[Eq 1]}$$

where

N = total number points to be sampled
p = expected percent accuracy
$q^{\sim}$ = $100 - p$
E = allowable error.

The equation above computes the ideal number of pixels to sample as reference points for an overall accuracy assessment of a classification. As allowable error increases, the number of required sample points decreases. Assuming a stratified random sampling approach, the total number of reference pixels or sample points required at a given expected accuracy and allowable error must be further stratified by thematic category. Van Genderen states that a minimum sample size of 20 is required for an 85 percent classification accuracy, while 30 observations (reference pixels) per class are required for 90 percent accuracy (at the 0.05 confidence level) (Van Genderen and Lock 1977).

Locating random points. The simplest way to generate random points is to pick two random numbers, one the horizontal and the other the vertical coordinate. In the UTM coordinate system, one random number would be chosen for the *easting* and another random number would be chosen for the *northing*. This is simple to do in a GIS. Using GRASS, the program *r.random* can be used to identify random pixels in a raster map (Westervelt et al. 1987).

Error matrix. An error matrix can be useful when evaluating the effectiveness of a discrete classification of remotely-sensed data. An error matrix is a means of reporting site-specific error (Campbell 1987). The error matrix is derived from a comparison of *reference* map pixels to the *classified* map pixels and is organized as a two dimensional matrix. This matrix takes the form of the columns representing the reference data by category and rows representing the classification by category. An error matrix is also referred to as a confusion matrix or contingency table, and in many cases, classification categories are arranged in columns and reference data represented along the rows of the matrix (Janssen and van der Well 1994). However, for consistency and ease of explanation, this document assumes an error matrix arranged according to the original definition shown in Table 1.

Table 1. Error matrix.

		Reference Data			
		Grass	Forest	Water	Row Marginals
Classified Data	Grass	77	8	0	85.00
	Forest	6	84	0	90.00
	Water	0	0	74	74.00
Column Marginals		83.00	92.00	74.00	249.00

Measures of agreement. From the error matrix, several measures of classification accuracy can be calculated, including percentage of pixels correctly classified, errors of omission, and errors of commission. In addition, statistical measures such as the Kappa Coefficient of Agreement, Kappa variance, and Kappa standard normal deviate can be calculated from the error matrix. The most commonly used measure of agreement is percentage of pixels correctly classified. This measure (Equation 2) is

$$\frac{\sum_{i=1}^{r} x_{ii}}{\sum_{i=1}^{r} x_{i+}} \qquad \text{[Eq 2]}$$

simply the number of pixels correctly classified from the validation set of pixels divided by the total number of reference pixels. Percentage correct is calculated by dividing the sum of the diagonal entries of the error matrix by the total number of reference pixels. Therefore, percent correct provides an overall accuracy assessment of a classification. However, if a minimum classification accuracy is required, it is necessary to verify that the calculated percent correct for the overall classification does indeed exceed the pre-determined minimum classification accuracy with some level of confidence. To assure that a minimum overall accuracy, a one-tailed lower confidence limit at a specific level of confidence must exceed the minimum accuracy standard (Jensen 1986). For example, the lower confidence limit for a one-tailed binomial distribution at a 95 percent confidence level can be calculated by Equation 3:

$$p = p^{\sim} - \left[1.645 \sqrt{(p^{\sim})(q^{\sim})/n} + \frac{50}{n} \right] \qquad \text{[Eq 3]}$$

where:

p = the accuracy of the map expressed as a percent

n = the sample size

$p^{\sim}$ = percent of observation correctly classified

$q^{\sim}$ = $100 - p^{\sim}$.

If p exceeds the minimum accuracy required of the classification, then the accuracy of the classification meets or exceeds the minimum accuracy requirement at the 95 percent confidence level. Percent correct provides an overall indication of how well a classification performed. However, an error matrix provides not only information that can be used to assess overall classification accuracy, but also information about the performance of a classification on a category-by-category basis.

To assess the classification accuracy of individual categories, the percent correct by category can be calculated. Percent correct ($p^{\sim}$) for an individual category is calculated by dividing the total number of correctly classified pixels for that category, i.e., the diagonal entry, by the total number of pixels in the reference map for that category, i.e., the column total (Table 1). As with the overall accuracy assessment, it is also necessary to determine if the accuracy of classification for individual categories exceeds some minimum accuracy requirement at some level of confidence. However, to determine the confidence limits of percentage correct for individual categories, a two-tailed test is appropriate. The upper and lower confidence limits are calculated in much the same way as the lower confidence limit for the overall percent correct is calculated, except that a two-tailed test is necessary. For example, the upper and lower confidence limits for a two-tailed binomial distribution at a 95 percent confidence level can be calculated using Equation 4 (Jensen 1986).

$$p = p \pm \left[1.96\sqrt{(p^{\sim})(q^{\sim})/n} + \frac{50}{n}\right] \qquad \text{[Eq 4]}$$

where:

p = the 95 percent confidence limits
$p^{\sim}$ = the percent correct for the category
$q^{\sim}$ = 100 - $p^{\sim}$
n = the number of observations in a particular category.

As with overall percentage correct calculations, if the confidence interval for percentage correct for an individual category is greater than the minimum required accuracy for a specific category, then the accuracy of classification of that individual category meets or exceeds the minimum accuracy for that category at a certain level of confidence. In addition to providing information necessary to calculate percentage correct for an overall classification or for individual categories with respective confidence intervals, an error matrix also contains other information useful in assessing the accuracy of a classification. The diagonal that extends from the upper left corner to the lower right corner of the matrix is referred to as "the diagonal," where each diagonal entry represents the number of correctly classified pixels for that specific category. Diagonal entries were used in the above examples to calculate percentage correct with respective confidence intervals. Assuming the arrangement

of the error matrix as discussed earlier with reference categories spread across the top (x-axis) of the matrix and classification categories distributed along the left hand side (y-axis) (Table 1), the column of sums on the right hand side represents the number of pixels in each category of the classified image under evaluation, and the bottom row of sums represents the total number of pixels in each category of the reference map. These sums are referred to as *row* and *column marginals*. In addition, nondiagonal values in each column represent *errors of omission* and nondiagonal values in each row represent *errors of commission* (Campbell 1987).

Errors of omission refer to pixels in the reference map that were classified as something other than their "known" or "accepted" category value. In other words, pixels of a known category were *excluded* from that category due to classification error. Errors of commission, on the other hand, refer to pixels in the classification map that were incorrectly classified and do not belong in the category in which they were assigned according to the classification. In other words, pixels in the classified image are *included* in categories in which they do not belong. Referring back to the error matrix, errors of omission for each category are computed by dividing the sum of incorrectly classified pixels in the nondiagonal entries of that category *column* by the total number of pixels in that category according to the reference map (i.e., the column marginal or column total). In a like manner, errors of commission for each category are calculated by dividing the sum of incorrectly classified pixels in the nondiagonal entries of that category *row* by the total number of pixels in that category according to the reference map (i.e., the column marginal or total) (Jensen 1986).

When evaluating the accuracy of an overall classification, it is best to examine several measures of accuracy, including overall percentage correct, percentage correct by category and also both errors of commission and omission by category. Examination of a single measure of accuracy may lead to incorrect assumptions about the accuracy of a classification. Different accuracy measures may be of interest depending on whether the person performing the classification is interested in the success of the classification or the end user of the classified map is interested in the accuracy or reliability of the map. A person interested in evaluating their classification efforts may be more interested in producer accuracy, which is simply the percentage of pixels of a know category type in a reference map that were actually classified as such. An end user of the map, however, may be more concerned with the reliability of the map, or user accuracy, which is simply the percentage of pixels in each category of the classification map that are actually that category of the ground (Congalton 1991). Obviously, both user and producer accuracy should be of interest and are important measures of accuracy.

User accuracy, or reliability, is actually the equivalent of percentage correct for an individual category and is calculated as described earlier. Producer accuracy is calculated in a similar fashion, with the only difference being that the total number of correctly classified pixels for a category is divided by the total number of pixels in that category in the *classification* map (i.e., the row marginal or row total) instead of dividing by the total number of pixels in that category in the *reference* map (i.e., the column marginal or column total). User and producer accuracy are directly related to errors of commission and errors of omission, respectively (Janssen and van der Wel 1994). The relationships are:

User's Accuracy (reliability) = percentage correct by category = 100% - error of commission (%)

and

Producer's Accuracy = 100% - error of omission (%)

Although percentage correct is the most commonly used measure of accuracy assessment, this measure has limitations. It is only suitable when making comparisons between another classification with the same resulting end number of categories. Commonly used measures of accuracy assessment such as percent correct, user accuracy, and producer accuracy are also limited by the fact that they do not account for simple random chance of assigning pixels to correct categories. Surprisingly enough, simple random assignment of pixels to categories could potentially lead to good results (Campbell 1987). Obviously, pixels are not assigned randomly during image classification, but there are statistical measures that attempt to account for the contribution of random chance when evaluating the accuracy of a classification. The Kappa Coefficient of Agreement is a statistic suitable for assessing accuracy of nominal data classification.

The Kappa Coefficient is a discrete multivariate measure that differs from the usual measures of overall accuracy assessment in basically two ways. First, the calculation takes into account all of the elements of the error matrix, not just the diagonals of the matrix (Foody 1992). This has the effect of taking into account chance agreement in the classification. The resulting Kappa measure compensates for chance agreement in the classification and provides a measure of how much better the classification performed in comparison to the probability of random assigning of pixels to their correct categories. Estimated variance of the Kappa Coefficient of Agreement can also be calculated. This is most useful in comparing two different approaches to the same classification scheme by allowing a standard normal deviate, or Z score, to be calculated. The Z score is used to determine if the differences in accuracy levels for two classifications with the same resultant classification scheme are significant. "The Kappa test statistic tests the null hypothesis that two independent classifiers do not agree on the rating or classification of the same physical object, in this case the class

of a ground truth site" (Fitzgerald and Lees 1994). The Kappa Coefficient of Agreement, $\hat{k}$, is calculated as:

$$\hat{k} = \frac{N \sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r} (x_{i+} * x_{i+})}{N^2 - \sum_{1=1}^{r} (x_{i+} * x_{+i})} \qquad \text{[Eq 5]}$$

where:

r = the number of rows in the error matrix
x_{ii} = the number of observations in row I and column i
x_{i+} = the marginal totals of row i
x_{+i} = the marginal totals of column i
N = the total number of observations (Bishop, Feinberg, and Holland 1975).

An advantage of using the Kappa Coefficient is the ability to compare two classifications and determine if the accuracy level between the two classifications is significantly different. The first step in determining significance is to calculate the variance of the Kappa Coefficient of Agreement. The estimated variance of Kappa can be calculated as:

$$\hat{V} = \frac{1}{N(1-p_c)^4} \left\{ \sum_{i=1}^{m} p_{ii} \left[(1-p_c)(p_{i+} + p_{+i})(1-p_o)\right]^2 1 + (1-p_o)^2 \right. \\ \left. \sum_{i=1}^{m} \sum_{j=1}^{m} p_{ij} (p_{i+} + p_{+j})^2 - (p_o p_c - 2p_c + p_o)^2 \right\} \qquad \text{[Eq 6]}$$

where:

$\hat{V}$ = the estimated variance of Kappa
N = the total number of observations
m = the number of categories
p_c = the proportion of observations that agree by chance
p_o = the proportion of observations correctly classified (Gong and Howarth 1992).

Using the Kappa Coefficient of Agreement, $\hat{k}$, and its estimated variance of Kappa, $\hat{V}$ for each classification, the standard normal deviate, Z, can then be calculated as (Congalton, Oderwald, and Mead 1983):

$$Z = \frac{(\hat{k}_1 - \hat{k}_2)}{[V(\hat{k}_1) + V(\hat{k}_2)]_{½}} \qquad \text{[Eq 7]}$$

where:

Z = the standard normal deviate

$\hat{k}_1$ = the Kappa Coefficient of Agreement for the first classification

$\hat{k}_2$ = the Kappa Coefficient of Agreement for the second classification

$V(\hat{k}_1)$ = the estimated variance of $\hat{k}_1$

$V(\hat{k}_2)$ = the variance of $\hat{k}_2$.

If Z exceeds 1.96, then the difference is significant at the 95 percent confidence level (Rosenfield and Fitzpatrick-Lins 1986). If Z exceeds 2.58, then the difference is significant at the 99 percent confidence level (Gong and Howarth 1992). If it is found that no significant difference exists, either classification can be used since they are essentially the same in terms of accuracy. Many geographic information/image-processing systems have the capability to calculate $\hat{k}$ and $\hat{V}$.*

It is also possible to calculate a measure of agreement, Conditional Kappa Coefficient of Agreement, for each individual class. The conditional Kappa is somewhat analogous to the overall Kappa Coefficient of Agreement measure except that a Conditional Kappa Coefficient of Agreement can be derived for each category of the classification. The Conditional Kappa Coefficient of Agreement measure is used to evaluate classification accuracies on a class-by-class basis. The Conditional Kappa Coefficient of Agreement (Bishop, Feinberg, and Holland 1975) is calculated by:

$$K_i = \frac{p_{ii} - p_{i+}p_{+i}}{p_{i+} - p_{i+}p_{+i}} \quad \text{[Eq 8]}$$

where:

K_i = Conditional Kappa Coefficient of Agreement for the *ith* category

P_{ii} = the number of correct observation for the *ith* category

P_{i+} = the *ith* row marginal

P_{+i} = the *ith* column marginal.

Conclusion

A rigorous assessment of the accuracy of a discrete classification of remotely-sensed data requires more than a simple calculation of percent of overall correct. The use of discrete multivariate statistical techniques enhances the accuracy assessment process. The use of an error matrix can help identify problems with a classification and can help improve classification by isolating misclassifications of pixels. This can help identify appropriate classification models and potential shortfalls in the type and

* The GIS/IP software GRASS has the ability to calculate these two values.

quality of ground-truth data. Site-specific accuracy assessment is essential when the resultant classification is used as an input for some model or as a basis for a management decision.

3 Accuracy Assessment Case Studies

This chapter presents two different classifications of a synthetic data set and their corresponding error matrices as examples of both accuracy assessment of a single classification and comparison of classification accuracies between two different classifications.

Single Classification Accuracy Assessment

Error matrix one summarizes the classification of a satellite image into four categories: Woodland, Grassland, Non-Vegetated, and Water. Reference pixels for each of the four categories were selected in a stratified random fashion. The number of reference pixels for each category are the column marginal in error matrix one. As previously mentioned, a *minimum* of 30 reference pixels per category is required to provide meaningful results at the 90 percent accuracy level with an allowable error of 5 percent. In addition, according to formula 1, the *ideal* total number of reference pixels for the same expected accuracy level and allowable error is 144. This example meets both criteria with a minimum of 48 reference pixels in the Grassland and Water categories, and a total of 200 reference pixels.

Classification–Matrix 1

The most common measure of overall accuracy is percent correct. In this example, 180 out of 200 pixels were correctly classified, resulting in 90 percent correct. Referring to the error matrix, percent correct was calculated by summing the diagonal values and dividing by the total number of pixels. Given a 90 percent overall observed correct, one could reach the initial conclusion that the classification performed well. However, if a level of classification accuracy for the product map is required prior to the classification, it is necessary to test the level of statistical confidence to the accuracy to ensure that the classification actually exceeds the minimum accuracy required with some level of confidence. In this example, assuming a predetermined minimum standard of 90 percent, it can be said with 95 percent confidence that the classification meets the 90 percent accuracy criteria according to Equation 3.

A more thorough review of the error matrix reveals additional information about the performance of the classification, including information about classification accuracy of individual categories (Tables 2 and 3). The classification performed best for the water category, with 48 out of 50 pixels classified correctly. The two remaining pixels that were classified as water were in fact nonvegetated surfaces, resulting in a 4 percent error of commission. Error of omission for water was zero percent, indicating that all 48 reference pixels in the water category were correctly classified. The classification was least successful in correctly classifying grassland areas, with only 40 of the 50 grassland pixels classified correctly. Examination of errors of commission and omission for grassland and nonvegetated surfaces indicates that the distinction between these two categories was the largest source of error or confusion in the classification. Six pixels classified as grassland were in fact nonvegetated surfaces and four pixels were classified as woodland. Five pixels of grassland in the reference data set were actually classified as nonvegetated and three pixels as woodland. Overall, error of commission and omission for grassland were the highest of the four categories, at 20 percent and 16.5 percent respectively.

Table 2. Error matrix 1.

		Reference Data				
		Woodland	**Grassland**	**Nonvegetated**	**Water**	**Row Marginals**
Data Classification	**Woodland**	47	3	0	0	50.00
	Grassland	4	40	6	0	50.00
	Nonvegetated	0	5	45	0	50.00
	Water	0	0	2	48	50.00
	Column Marginals	51.00	48.00	53.00	48.00	200.00

Table 3. Summary of error matrix 1.

Category	**% Commission**	**% Omission**	**Estimated Kappa**
Woodland	6.000000	7.843137	0.919463
Grassland	20.000000	16.666667	0.736842
No Veg.	10.000000	15.094340	0.863946
Water	4.000000	0.000000	0.947368
Kappa	**Kappa Variance**		
0.866667	0.018000		
Observed Correct	**Total Observed**	**% Observed Correct**	
180	200	90.000000	

In addition to errors of commission and omission, estimated kappa values for the individual categories also provides some indication of individual category classification accuracies. Again, classification of water was most accurate, with an estimated kappa of 0.947, while grassland was the least accurate, with an estimated kappa of 0.736. Kappa values for both the entire classification and each individual category account for the contribution of random chance in the classification. For example, assignment of pixels to the water category was 94 percent more accurate than what could be expected from a random assignment of pixels to one of the four categories.

Classification–Matrix 2

The second error matrix (Tables 4 and 5) represents a classification of the same area using a different classification technique. Similar to the previous classification (Tables 2 and 3), the classification of water was most accurate, with the highest estimated kappa and lowest errors of commission and omission. Classification of grassland was again the least accurate, with the lowest estimated kappa value and the highest errors of commission and omission. However, further examination of the error matrix reveals that this classification scheme resulted in more confusion between woodland and grassland than the previous classification. The woodland category had the second highest errors of commission and omission and the second lowest estimated kappa. In this case, four pixels classified as woodland were in fact grassland, and one additional pixel was nonvegetated. Also, eight woodland pixels in the reference data set were incorrectly classified as grassland.

In this example, 174 out of 200 pixels were correctly classified, resulting in 87 percent correct. Referring to the error matrix, percent correct was calculated by summing the diagonal values and dividing by the total number of pixels. Given an 87 percent overall observed correct, one could reach the initial conclusion that this classification also performed well. However, if a level of classification accuracy for the product map is required prior to the classification, it is necessary to test the level of statistical confidence to ensure that the classification actually exceeds the minimum level of

Table 4. Error matrix 2.

		Reference Data				
		Woodland	Grassland	Nonvegetated	Water	Row Marginals
Data Classification	Woodland	45	4	1	0	50.00
	Grassland	8	36	5	1	50.00
	Nonvegetated	0	4	46	0	50.00
	Water	0	1	2	47	50.00
	Column Marginals	53.00	45.00	54.00	48.00	174.00

Table 5. Summary of error matrix 2.

Category	% Commission	% Omission	Estimated Kappa
Woodland	10.000000	15.094340	0.863946
Grassland	28.000000	20.000000	0.638710
Nonvegetated	8.000000	14.814815	0.890411
Water	6.000000	2.083333	0.921053
Kappa	**Kappa Variance**		
0.826667	0.017400		
Observed Correct	**Total Observed**	**% Observed Correct**	
174	200	87.000000	

confidence. In this example, assuming a predetermined minimum standard of 87 percent, it can be said with 95 percent confidence that this classification meets the 87 percent accuracy criteria according to Equation 3.

As in the first example, the error matrix reveals additional information about the performance of the classification (Tables 4 and 5). The classification performed best for the water category, with 47 out of 50 pixels classified correctly. Of the three remaining pixels that were classified as water, two were classified as nonvegetated surfaces and one as grassland, resulting in a 6 percent error of commission. Error of omission for water was 2 percent, indicating that 47 of 48 reference pixels in the water category were correctly classified. The classification was least successful in correctly classifying grassland areas, with only 36 of the 50 grassland pixels classified correctly. Examination of errors of commission and omission for grassland and nonvegetated surfaces indicates that the distinction between these two categories was the again the largest source of error or confusion in the classification. Five pixels classified as grassland were in fact nonvegetated surface, one pixel classified as grassland was water, and eight pixels were classified as woodland. Four pixels of grassland in the reference data set were actually classified as nonvegetated, four pixels as woodland, and one pixel as water. Overall, errors of commission and omission for grassland were the highest of the four categories, at 28 and 20 percent respectively.

In addition to errors of commission and omission, estimated kappa values for the individual categories also provide some indication of individual category classification accuracies. Again, classification of water was most accurate, with an estimated kappa of 0.921, while grassland was the least accurate, with an estimated kappa of 0.639. Kappa values for both the entire classification and each individual category account for the contribution of random chance in the classification. Assignment of pixels to the

water category was 92 percent more accurate than what could be expected from a random assignment of pixels to one of the four categories.

Multiple Classification Accuracy Assessment

An overall comparison of two different classification schemes is sometimes desired, especially if the producer of the classification is testing different classification methods and is interested in the relative increase or decrease in accuracy for each new classification which is tested. In the first example, the classification (Tables 2 and 3) resulted in a higher percent correct (90 vs. 87 percent) and a higher kappa value (0.867 vs. 0.827) than classification two (Tables 3 and 4). From these observations alone, one might conclude that classification method one produced more accurate results. One might also conclude that classification one had the most difficulty in distinguishing between grassland and nonvegetated surfaces, while classification two not only had problems distinguishing between grassland and nonvegetated surfaces, but also had problems distinguishing between woodland and grassland. However, one additional test can be conducted that tests if one classification is significantly different from another by calculating the standard normal deviant, Z, using kappa and estimated kappa variance according to formula 7. If the standard normal deviate exceeds 1.96, then the difference between the accuracy of the classifications is significant at the 95 percent confidence level. Likewise, if Z exceeds 2.58, then the difference is significant at the 99 percent confidence level. In the example used here, the standard normal deviant is 0.13; therefore, one cannot say that the accuracy of classifications one and two is significantly different at the 95 percent confidence level.

4 Conclusions

Once the classified image is integrated into a GIS, thereby becoming an information source for natural resource managers, accuracy assessment should become an integral part of any classification process. Accuracy assessment may include both non-site-specific and site-specific analyses. Non-site-specific analysis includes relatively simple comparisons of areal coverage of categories, while site-specific analyses range from simple percent-correct calculations to more complex multivariate statistical techniques.

This study compared methods of site-specific and non-site-specific analysis, and concludes that non-site-specific analysis has limited utility and is only useful for detecting gross errors in a classification. Site-specific analysis, however, provides critical information about the locational accuracy of a classification, and is therefore more rigorous and useful for use with such programs as LCTA. Assuming an adequate sampling method, measures of agreement such as overall percent correct and percent correct for individual categories can be assigned statistically-defined confidence intervals to ensure that classifications meet minimum accuracy requirements.

Use of an error matrix is one method that can provide additional insight into classification errors that may be unique to specific categories, and may generate information necessary to calculate errors of commission and omission.

Use of the Kappa Coefficient of Agreement accounts for random chance in accuracy assessment. Kappa and estimated Kappa variance for two different classifications can also be used to calculate a standard normal deviate, which in turn can be used to determine if the kappa values of the two classifications are significantly different. Conditional Kappa Coefficients of Agreement can also be calculated to assess accuracies on a category-by-category basis.

Time and funding constraints may often dictate the amount of data that can be gathered in conjunction with the collection of satellite imagery; a comprehensive accuracy assessment may not always be practical. However, depending on the intended use of the classified data, some level of accuracy assessment should always be performed. It is concluded that, at a minimum, a Kappa Coefficient of Agreement should be attached to any resultant classification of satellite imagery. Ideally, several

measure-of-accuracy assessments should be performed and included as documentation with the classification. This is critical to end-users of the data, and also provides valuable metadata that will be necessary as more stringent standards are imposed on the exchange of digital spatial data between end-users.

Bibliography

Bishop, Yvonne M.M., Stephen E. Feinberg, and Paul W. Holland, *Discrete Multivariate Analysis* (MIT Press, Cambridge, MA, 1975), p 396.

Campbell, James B., *Introduction to Remote Sensing* (The Guilford Press, New York, 1987), p 340.

Congalton, Russel G., "A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data," *Remote Sensing of Environment*, vol 37 (1991), pp 35-46.

Congalton, R.G., R.G. Oderwald, and R.A. Mead, "Assessing LANDSAT Classification Accuracy Using Discrete Multivariate Statistical Techniques," *Photogrammetric Engineering and Remote Sensing*, vol 49, No. 2 (1983), pp 1671-1678.

Engineer Technical Note (ETN) 420-74-3, *Army Land Inventory and Monitoring Procedures on Military Installations* (U.S. Army Engineering and Housing Support Center [USAEHSC], Fort Belvoir, VA, 1990).

Fitzgerald, R.W., and B.G. Lees, "Assessing the Classification of Multisource Remote Sensing Data," *Remote Sensing of Environment*, vol 47 (1994), pp 362-368.

Foody, Giles M., "On the Compensation for Chance Agreement in Image Classification Accuracy Assessment," *Photogrammetric Engineering and Remote Sensing*, vol 58, No. 10 (1992), pp 1459-1460.

Green, K., "Spatial Imagery and GIS," *Journal of Forestry*, vol 90, No. 11 (1992), pp 32-36.

Gong, Peng, and Philip J. Howarth, "Frequency-Based Contextual Classification and Gray-Level Vector Reduction for Land-Use Identification," *Photogrammetric Engineering and Remote Sensing*, vol 58, No. 4 (1992), p 425.

Hord, R. Michael, and William Brooner, "Land-Use Map Accuracy Criteria," *Photogrammetric Engineering and Remote Sensing*, vol 42, No. 5 (1976), pp 671-677.

Janssen, Lucas L.F., Frans J.M. van der Wel, "Accuracy Assessment of Satellite Derived Land-Cover Data: A Review," *Photogrammetric Engineering and Remote Sensing*, vol 60, No. 4 (1994), pp 419-426.

Jensen, John R., *Introductory Digital Image Processing* (Prentice-Hall, Englewood Cliffs, NJ, 1986), p 227.

Lillesand, T.M., and R.W. Kiefer, *Remote Sensing and Image Interpretation* (John Wiley & Sons, New York, 1987).

Rosenfield, George H., and Katherine Fitzpatrick-Lins, "A Coefficient of Agreement as a Measure of Thematic Classification Accuracy," *Photogrammetric Engineering and Remote Sensing*, vol 52, No. 2 (1986), pp 223-227.

Shapiro, M., C. Bouman, and C.F. Bagley, *A Multiscale Random Field Model for Bayesian Image Segmentation*, Technical Report (TR) EC-94/21/ADA283875 (U.S. Army Construction Engineering Research Laboratory, [USACERL], 1994).

Stehman, Stephen V., "Comparison of Systematic and Random Sampling for Estimating the Accuracy of Maps Generated From Remotely-Sensed Data," *Photogrammetric Engineering & Remote Sensing*, vol 58, No. 9 (1992), pp 1343-1350.

Tazik, David J., Steven D. Warren, Victor E. Diersing, Robert B. Shaw, Robert J. Brozka, Calvin F. Bagley, and William R. Whitworth, *U.S. Army Land Condition-Trend Analysis (LCTA) Plot Inventory Field Methods*, TR N-92/03/ADA247931 (USACERL, February 1992).

Van Genderen, J.L., and B.F. Lock, "Testing Land-Use Map Accuracy," *Photogrammetric Engineering and Remote Sensing*, vol 43, No. 9 (1976), pp 1135-1137.

Westervelt, James D., Michael Shapiro, William D. Goran, and David P. Gerdes, *Geographic Resources Analysis Support System (GRASS) Version 4.0 User's Reference Manual*, Automated Data Processing (ADP) Report N-87/22/ADA255218 (USACERL, June 1992).

USACERL Distribution

Chief of Engineers
ATTN: CECHEC-IM-LH (2)
ATTN: CECHEC-IM-LP (2)
ATTN: CERD-L
ATTN: DAEN-ZCI-P (2)
ATTN: CECW-PF
ATTN: CECW-EP-S
ATTN: CECW-RE
ATTN: CEIM-P
ATTN: CECW-OR
ATTN: CECC-R

US Army Engineer District
ATTN: Chief, Regulatory Functions
Buffalo 14207
Norfolk 23510
Huntington 25701
Wilmington 28402
Charleston 29402
St. Paul 55101
Chicago 60606
Little Rock 72203
Galveston 77553
Albuquerque 87103
Los Angeles 90012
San Francisco 94105
Sacramento 95814
Portland 97208
Seattle 99362
New England 02254
ATTN: CENED-OD-R
New York 10278
ATTN: CENAN
Philadelphia 19107
ATTN: CENAP
Baltimore 21203
ATTN: CENAB-OP-R
Savannah 31402
ATTN: CESAS-OP-FP
Jacksonville 32232
ATTN: CESAJ-R-RD
Mobile 36628
ATTN: CESAM-OP-S
Nashville 37202
ATTN: CEORN
Memphis 38103
ATTN: CELMM-CO-R
Vicksburg 39180
ATTN: CELMV
Louisville 40201
ATTN: CEORL
Detroit 48231
ATTN: CENCE-CO
Rock Island 61204
ATTN: CENCR
St. Louis 63101
ATTN: CELMS-OD-F
Kansas City 64106
ATTN: Chief, Permits Section
Omaha 68102
ATTN: CEMRO-OP
New Orleans 71060
ATTN: Chief, Permits Section
Tulsa 74121
ATTN: CEWST
Fort Worth 76102
ATTN: CESWF-OD-O

US Army Engr Divisions
North Central 60606
ATTN: CENCD-CO
Southwest 75242
ATTN: CESWD-CO-R

US Military Academy 10996
ATTN: Dept of Geo & Envr Engr (2)
ATTN: Natural Resources Branch

Aberdeen Proving Ground, MD 21005
ATTN: ISC-TECOM
ATTN: HSHB-CI

US EPA Research Lab 97333

Yakima Firing Center 98901

CECPW 22060
ATTN: CECPW-FN
ATTN: CECPW-SP
ATTN: CECPW-FM-A

Bureau of Land Management
WASH DC 20240
Fort Collins 80526
Denver 80225

US Dept of Commerce 20233

Army National Guard 22204
ATTN: NGB-AREC

US Army Concepts Analysis Agency 20814

FBI Academy 22135

USA Foreign Science Tech Ctr 22901

Naval Oceanographic Office 39522

Redstone Arsenal 35898
ATTN: AMSMI-RA-EH-MP-PC

CEWES
ATTN: CEWES-IM-DA 39181
ATTN: CADD Center 39180

Wright-Patterson AFB 45433
ATTN: NSBIT AL/OEBN

Michigan Dept of Military Affairs 48913

Twin Cities Army Ammo Plant (2) 55112

Camp Ripley 56345
ATTN: Ofc of Archeology & Engr (2)

5th Inf. Fort Polk 71459
ATTN: Envr & Nat Rsrcs Mgmt Div (2)

US Army Materiel Cmd 61299
ATTN: AMXEN-U (2)

US Army Europe
HQ USAREUR 09403
ATTN: AEAEN-FE-E (2)
V Corps 09079
ATTN: AETV-EHF-R

Texas Army Nat'l Guard 78763

Lone Star Army Ammo Plant 75505

Red River Army Depot
ATTN SDSRR-GB 75507

Dugway Proving Ground 84022
ATTN: DPG-EN-E (2)

Yuma Proving Ground 85365
ATTN: ATEYP-ES-E

White Sands Missile Range
ATTN: STEWS-ES-E 88002

Envr Response & Info Ctr
ATTN: ENVR-EP 20310

Nat'l Geophysical Data Ctr
ATTN: Code E-GCI 80303

Hohenfels Training Area 09173
ATTN: AETTH-DEH
ATTN: AETTH-DEH-ENV-APO

US Army Forts
Fort Belvoir, VA 22060
ATTN: CETEC-CA-D
ATTN: AMSEL-RD-NV-VMD-TST
ATTN: Envr & Nat Res Div
Fort Monroe, VA 23651
ATTN: ATBO-GE
Fort Drum 13603
ATTN: AFZS-EH-E
Fort Jackson 29207
ATTN: ATCJ-EHN
Fort Gillem 30050
ATTN: FCEN-CED-E
Fort Gordon 30905
ATTN: ATZH-DIE (2)
Fort Stewart 31314
ATTN: AFZP-DEN-W
Fort Benning 31905
ATTN: Nat. Resource Mgmt Div (2)
Fort McClellan 36205
ATTN: ATZN-FEE
Fort Rucker 36362
ATTN: ATZQ-EH
Fort Knox 40121
ATTN: ATZK-EHE
Fort Campbell 42223
ATTN: AFZH-DEH
Fort Benjamin Harrison 46216
ATTN: ATZI-ISP (2)
Fort McCoy 54656
ATTN: AFZR-DEN
Fort Riley 66442
ATTN: AFZN-DE-N (2)
Fort Chaffee 72905
ATTN: ATZR-ZFE (2)
Fort Sill 73503
ATTN: Fish & Wildlife Br (2)
Fort Leonard Wood 65473
ATTN: ATZT-DEH-EE
Fort Dix 08640
ATTN: ATZD-EHN
Fort Eustis 23604
ATTN: Ranges & Targets Dir
Fort Worth 76115
ATTN: Cartographic Ctr (2)
Fort Hood 76544
ATTN: AFZF-DE-ENV
Fort Bliss 79916
ATTN: ATZC-DEH-E
Fort Carson 80913
ATTN: AFZC-ECM-NR
Fort Huachuca 85613
ATTN: ATZS-EHB
Fort Irwin 92310
ATTN: AFZJ-EH
Fort Lewis 98433
ATTN: AFZH-DEQ
ATTN: ATZH-EHQ
Fort Richardson 99505
ATTN: DPW
Fort Bragg 28307
ATTN: DPW

National Weather Service 20910

US Geological Survey 22092

Pine Bluff Arsenal 71602
ATTN: SMCPB-EMB

US Army Topographic Engr Center 22060
ATTN: CETEC-IM-T

US Army Cold Regions Res & Engr Lab
ATTN: CECRL-IS 03755

NASA/SSC/STL 39529

Defense Technical Info Center
ATTN: DTIC-FAB (2)

141
4/95